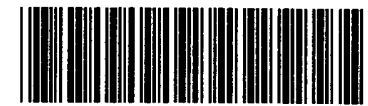


0590  
07/8



OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/991,888

DATE: 07/25/2002  
TIME: 14:55:38

Input Set : N:\Crf3\RULE60\09991888.raw  
Output Set: N:\CRF3\07252002\I991888.raw

SEQUENCE LISTING

- 1 (1) GENERAL INFORMATION:
- 2 (i) APPLICANT: LI, SHENGFENG
- 3 PHILLIPS, DAVID
- 4 (ii) TITLE OF INVENTION: IDENTIFICATION OF Bap-1, A PROTEIN THAT
- 5 BINDS TO INTEGRIN AND IS INVOLVED IN INTEGRIN-
- 6 MEDIATED
- 7 SIGNAL TRANSDUCTION
- 8 (iii) NUMBER OF SEQUENCES: 5
- 9 (iv) CORRESPONDENCE ADDRESS:
- 10 (A) ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
- 11 (B) STREET: 1800 M ST., NW
- 12 (C) CITY: WASHINGTON
- 13 (D) STATE: DC
- 14 (E) COUNTRY: USA
- 15 (F) ZIP: 20036
- 16 (v) COMPUTER READABLE FORM:
- 17 (A) MEDIUM TYPE: Floppy disk
- 18 (B) COMPUTER: IBM PC compatible
- 19 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- 20 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- 21 (vi) CURRENT APPLICATION DATA:
- C--> 22 (A) APPLICATION NUMBER: US/09/991,888
- C--> 23 (B) FILING DATE: 26-Nov-2001
- 24 (vii) PRIOR APPLICATION DATA:
- 25 (A) APPLICATION NUMBER: US/09/590,175
- 26 (B) FILING DATE: 09-Jun-2000
- 27 (A) APPLICATION NUMBER: US 08/753,038
- 28 (B) FILING DATE: 1996-11-18
- 29 (A) APPLICATION NUMBER: US 08/972,719
- 30 (B) FILING DATE: 1997-11-18
- 31 (viii) ATTORNEY/AGENT INFORMATION:
- 32 (A) NAME: REID G. ADLER
- 33 (B) REGISTRATION NUMBER: 30,988
- 34 (C) REFERENCE/DOCKET NUMBER: 44481-5025-02-US
- 35 (ix) TELECOMMUNICATION INFORMATION:
- 36 (A) TELEPHONE: (202) 467-7000
- 37 (B) TELEFAX: (202) 467-7258
- 38 (2) INFORMATION FOR SEQ ID NO: 1:
- 39 (i) SEQUENCE CHARACTERISTICS:
- 40 (A) LENGTH: 3467 base pairs
- 41 (B) TYPE: nucleic acid
- 42 (C) STRANDEDNESS: double

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/991,888

DATE: 07/25/2002

TIME: 14:55:38

Input Set : N:\Crf3\RULE60\09991888.raw

Output Set : N:\CRF3\07252002\I991888.raw

```

43      (D) TOPOLOGY: linear
44      (ii) MOLECULE TYPE: DNA (genomic)
45      (ix) FEATURE:
46          (A) NAME/KEY: CDS
47          (B) LOCATION: 226..1233
48      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
49      GCCCCTCGCT CGCTCGCTCC TTCCCGCCCT CCCCAGCAGC CCGGCCGAGC CGGCTTCCCC      60
50      TCAGTCTCTC ATGAATATTG AGCGGCCCTT GTTGTATTTT CCGAGCTCCA TTGCGGAAGC      120
51      TGAGGCTCGC CATATTGTGC GCGGCGCCCG GCGTCCGCGG CAGCTGATAC CAGAGTCTTG      180
52      CTCCGCGCCG GGCAGCGGA GCCCTGGGCT GGGGAGGAG CCGCA ATG TCT CAG      234
53                                     Met Ser Gln
54                                     1
55      GCT GTG CAG ACA AAC GGA ACT CAA CCA TTA AGC AAA ACA TGG GAA CTC      282
56      Ala Val Gln Thr Asn Gly Thr Gln Pro Leu Ser Lys Thr Trp Glu Leu
57          5                      10                      15
58      AGT TTA TAT GAG TTA CAA CGA ACA CCT CAG GAG GCA ATA ACA GAT GGC      330
59      Ser Leu Tyr Glu Leu Gln Arg Thr Pro Gln Glu Ala Ile Thr Asp Gly
60      20                      25                      30                      35
61      TTA GAA ATT GTG GTT TCA CCT CGA AGT CTA CAC AGT GAA TTA ATG TGC      378
62      Leu Glu Ile Val Val Ser Pro Arg Ser Leu His Ser Glu Leu Met Cys
63          40                      45                      50
64      CCA ATT TGT TTG GAT ATG TTG AAG AAC ACC ATG ACT ACA AAG GAG TGT      426
65      Pro Ile Cys Leu Asp Met Leu Lys Asn Thr Met Thr Thr Lys Glu Cys
66          55                      60                      65
67      TTA CAT CGT TTT TGT GCA GAC TGC ATC ATC ACA GCC CTT AGA AGT GGC      474
68      Leu His Arg Phe Cys Ala Asp Cys Ile Ile Thr Ala Leu Arg Ser Gly
69          70                      75                      80
70      AAC AAA GAA TGT CCT ACC TGT CGG AAA AAA CTA GTT TCC AAA AGA TCA      522
71      Asn Lys Glu Cys Pro Thr Cys Arg Lys Lys Leu Val Ser Lys Arg Ser
72      85                      90                      95
73      CTA AGG CCA GAC CCA AAC TTT GAT GCA CTC ATC AGC AAA ATT TAT CCA      570
74      Leu Arg Pro Asp Pro Asn Phe Asp Ala Leu Ile Ser Lys Ile Tyr Pro
75      100                      105                      110                      115
76      AGT CGT GAT GAG TAT GAA GCT CAT CAA GAG AGA GTA TTA GCC AGG ATC      618
77      Ser Arg Asp Glu Tyr Glu Ala His Gln Glu Arg Val Leu Ala Arg Ile
78          120                      125                      130
79      AAC AAG CAC AAT AAT CAG CAA GCA CTC AGT CAC AGC ATT GAG GAA GGA      666
80      Asn Lys His Asn Asn Gln Gln Ala Leu Ser His Ser Ile Glu Glu Gly
81          135                      140                      145
82      CTG AAG ATA CAG GCC ATG AAC AGA CTG CAG CGA GGC AAG AAA CAA CAG      714
83      Leu Lys Ile Gln Ala Met Asn Arg Leu Gln Arg Gly Lys Lys Gln Gln
84          150                      155                      160
85      ATT GAA AAT GGT AGT GGA GCA GAA GAT AAT GGT GAC AGT TCA CAC TGC      762
86      Ile Glu Asn Gly Ser Gly Ala Glu Asp Asn Gly Asp Ser Ser His Cys
87          165                      170                      175
88      AGT AAT GCA TCC ACA CAT AGC AAT CAG GAA GCA GGC CCT AGT AAC AAA      810
89      Ser Asn Ala Ser Thr His Ser Asn Gln Glu Ala Gly Pro Ser Asn Lys
90      180                      185                      190                      195
91      CGG ACC AAA ACA TCT GAT GAT TCT GGG CTA GAG CTT GAT AAT AAC AAT      858

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/991,888

DATE: 07/25/2002

TIME: 14:55:38

Input Set : N:\Crif3\RULE60\09991888.raw

Output Set: N:\CRF3\07252002\I991888.raw

92	Arg Thr Lys Thr Ser Asp Asp Ser Gly Leu Glu Leu Asp Asn Asn Asn	
93	200 205 210	
94	GCA GCA ATG GCA ATT GAT CCA GTA ATG GAT GGT GCT AGT GAA ATT GAA	906
95	Ala Ala Met Ala Ile Asp Pro Val Met Asp Gly Ala Ser Glu Ile Glu	
96	215 220 225	
97	TTA GTA TTC AGG CCT CAT CCC ACA CTT ATG GAA AAA GAT GAC AGT GCA	954
98	Leu Val Phe Arg Pro His Pro Thr Leu Met Glu Lys Asp Asp Ser Ala	
99	230 235 240	
100	CAG ACG AGA TAC ATA AAG ACT TCT GGT AAC GCC ACT GTT GAT CAC TTA	1002
101	Gln Thr Arg Tyr Ile Lys Thr Ser Gly Asn Ala Thr Val Asp His Leu	
102	245 250 255	
103	TCC AAG TAT CTG GCT GTG AGG TTA GCT TTA GAA GAA CTT CGA AGC AAA	1050
104	Ser Lys Tyr Leu Ala Val Arg Leu Ala Leu Glu Glu Leu Arg Ser Lys	
105	260 265 270 275	
106	GGT GAA TCA AAC CAG ATG AAC CTT GAT ACA GCC AGT GAG AAG CAG TAT	1098
107	Gly Glu Ser Asn Gln Met Asn Leu Asp Thr Ala Ser Glu Lys Gln Tyr	
108	280 285 290	
109	ACC ATT TAT ATA GCA ACA GCC AGT GGC CAG TTC ACT GTA TTA AAT GGC	1146
110	Thr Ile Tyr Ile Ala Thr Ala Ser Gly Gln Phe Thr Val Leu Asn Gly	
111	295 300 305	
112	TCT TTT TCT TTG GAA TTG GTC AGT GAG AAA TAC TGG AAA GTG AAC AAA	1194
113	Ser Phe Ser Leu Glu Leu Val Ser Glu Lys Tyr Trp Lys Val Asn Lys	
114	310 315 320	
115	CCC ATG GAA CTT TAT TAC GCA CCT ACA AAG GAG CAC AAA TGAGCCTTTA	1243
116	Pro Met Glu Leu Tyr Tyr Ala Pro Thr Lys Glu His Lys	
117	325 330 335	
118	AAAACCAATT CTGAGACTGA ACTTTTTTAT AGCCTATTTT TTTAATATTA AAGATGTACT	1303
119	GGCATTACTT TTATGGAGAT CTTGGATATG TTGTTCAATT TTCTTTCTGA GCCAGACTAG	1363
120	TTTACGCTAT TCAAATCTTT TCCCCTTTAT TTAAGATTTT CTTTTTGGA GGGACTGCAA	1423
121	TTATTCAGTA TTTTTTCTT TCCTTTAAAA AAATATATCT GAAGTTTCTT GTGTTTTTTT	1483
122	TTTTCCCCAC AAAGTGTGTT TCCACTTGGA GCACCATTTT GACCCAGGAA TTTTTCATAG	1543
123	TTTCTGTATT CTTATAAGAT TCAGTTGGCT GTCCTTTTCC TGCTCCCTC AAAAGATTTT	1603
124	TAGTCATACA GAATGTTAAA TATTATGTAT TCTGACTTTT TTTTCCCCC GGAGTCTTGT	1663
125	ATATTTATAG TTTTCCTATA TAAACTGTAG TATCTTCATG AAGAACCCAA GGCTCAAATT	1723
126	TACTGTCCTT AAAAACAATT CTCATAGGAT TATTCTTTT ATGGTATTTT CTTCCATAAT	1783
127	ATCTCATTTT AAAAAGAAGT TCTTTATGAA ACTTAGTGTC CATTGTCATG CAATGTTTTT	1843
128	TTTTTCCATT CTTTTTCCCC TGTAATTTTG GAATTTCTGG TCCTGGGAAG AATCAAACAA	1903
129	AATCTTAAGT TCTATGAGAA CTTGGTTCAT TGACATATTC TGCTGAAGAA AGAAAAATTA	1963
130	AATTGGTAGT AAAATATAGT CTTCAAGTAT ACGTTTGAGA GTGCTTTTTT TTGTATTAGT	2023
131	TCTGCTGTCA CTTCAATTTCC TGTATTATAT GTGATGTTTT TCCCCATTAA AATACCAGAG	2083
132	ATAATGGAGA TATTTTGCAC TTTAGCCTTG ATGAAAAAGTA CAAGATATGT TCAAAGCTTC	2143
133	CCTAATTTTT TTCTTATTTG TAGCCACATA AGTTTCAAGA ATAACATGGC ACACAGAACA	2203
134	ATGGAAAAAA GTTTGTTTCC ATTGGAAGAT TATATCATTT TGGGTTGCCA CATCAGTTTA	2263
135	TAAATTTGGC GCTCTTTTAA TTACACTCTG TAGAAGGTTA ATAGAGCTTG AGCCCTGCTT	2323
136	TAATATGTAG TGAAAGATAA TTCTGTAGAA AAACGTCAGC CAGTAGGGTA AAGTCATTCT	2383
137	ACTGTTCTTA ATTTTTATAT TGAGGAACAA TATTGGGTGT TTGGGAGCCA GAAAGCTTTG	2443
138	TTGACAGATC AGAAATAAGA TTGACTTGGG TGTTATATTT CATCTCTCTC CAGACTCTAG	2503
139	GTATATTTCC AACTTTATAT ATCACAGTAT TTA AAAAGAC ATGTTTGCAT TGAGAAATTA	2563
140	ACCCTAAAGG GTTTTCAATA GGGTGTAGAC CTCCAGTACC TTTGTAAC TAAGTCTGTCT	2623

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/991,888

DATE: 07/25/2002

TIME: 14:55:38

Input Set : N:\Crif3\RULE60\09991888.raw

Output Set: N:\CRF3\07252002\I991888.raw

```

141 AGTCATTGTA AATATTTATC TGTCAGTTTT GACAGATTGG GGCCAGCTTG ATGTTTTTAAA 2683
142 TCTTCAGCCC GGTATGAAAA CTAAAGGTA TATATTCAAT TTTTACCAT TTTATGGAAA 2743
143 ATATTTAAAA TTTGTTTTTA CAGGGTTTTT TTTTTTTTTT TTTTTTTTTT GTAATCTGTG 2803
144 CCATGAAATT TGAAAACCAC CAAAAATCAA GGGAACTTTT ATATATTCAA TTCCTTTTCT 2863
145 GGTGTAATGT TAAAGTTGTA TAGATTATTA ATGCATGCCC ACTGAATATA ACCCTGGTTT 2923
146 TGTGATAAAA CTGCTTAGAT TTTGTTGATG ACATTAGATT AGTAGTTGCA TTAAATAACT 2983
147 AAATTCCTCAT TGTGATTAAT TGAAATTTTG TCTTTAAGCA GAGAGTTATT TGTGACTATA 3043
148 AGCTTTGTGC TTAGAGAATG TATGTGTTTT TATCTGTCAG TATGGGAGGA TATAAACTGC 3103
149 ATCATTAGTG AAATTATTGG TTGTGTAATC CTTTGTGAAA TATAATTCTA GGTATTTGAT 3163
150 AGGGTATTGA GTGTATTTTG TGTGTGTGTG GATGTGTGTT TTGGGGTACG GGGAGAGGCG 3223
151 ATGCTATTGG CCATCACTAC CAACCAGGGT TTCAAAAAGT ATATACCTAA GTAATTTCTT 3283
152 TTATCACTAC CTCAACTGAG GAAGAAAAGG CTCACCACAA GTGGTGTGAA GGCTTTGGGT 3343
153 ACTTAGTTCT AAATTTTTTT ATGGTAACAT ATACATAGCC ACATTTACAG TTTTAACCAT 3403
154 TTTAAGGCAT GTAATTCAGT GGGGTTAGGT ACATTACAA TGTTGTGTAA TGATCACCGC 3463
155 CGTG 3467

```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 336 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

164 Met Ser Gln Ala Val Gln Thr Asn Gly Thr Gln Pro Leu Ser Lys Thr
165 1 5 10 15
166 Trp Glu Leu Ser Leu Tyr Glu Leu Gln Arg Thr Pro Gln Glu Ala Ile
167 20 25 30
168 Thr Asp Gly Leu Glu Ile Val Val Ser Pro Arg Ser Leu His Ser Glu
169 35 40 45
170 Leu Met Cys Pro Ile Cys Leu Asp Met Leu Lys Asn Thr Met Thr Thr
171 50 55 60
172 Lys Glu Cys Leu His Arg Phe Cys Ala Asp Cys Ile Ile Thr Ala Leu
173 65 70 75 80
174 Arg Ser Gly Asn Lys Glu Cys Pro Thr Cys Arg Lys Lys Leu Val Ser
175 85 90 95
176 Lys Arg Ser Leu Arg Pro Asp Pro Asn Phe Asp Ala Leu Ile Ser Lys
177 100 105 110
178 Ile Tyr Pro Ser Arg Asp Glu Tyr Glu Ala His Gln Glu Arg Val Leu
179 115 120 125
180 Ala Arg Ile Asn Lys His Asn Asn Gln Gln Ala Leu Ser His Ser Ile
181 130 135 140
182 Glu Glu Gly Leu Lys Ile Gln Ala Met Asn Arg Leu Gln Arg Gly Lys
183 145 150 155 160
184 Lys Gln Gln Ile Glu Asn Gly Ser Gly Ala Glu Asp Asn Gly Asp Ser
185 165 170 175
186 Ser His Cys Ser Asn Ala Ser Thr His Ser Asn Gln Glu Ala Gly Pro
187 180 185 190
188 Ser Asn Lys Arg Thr Lys Thr Ser Asp Asp Ser Gly Leu Glu Leu Asp
189 195 200 205
190 Asn Asn Asn Ala Ala Met Ala Ile Asp Pro Val Met Asp Gly Ala Ser

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/991,888

DATE: 07/25/2002

TIME: 14:55:38

Input Set : N:\Crf3\RULE60\09991888.raw

Output Set: N:\CRF3\07252002\I991888.raw

```

191          210          215          220
192    Glu Ile Glu Leu Val Phe Arg Pro His Pro Thr Leu Met Glu Lys Asp
193    225          230          235          240
194    Asp Ser Ala Gln Thr Arg Tyr Ile Lys Thr Ser Gly Asn Ala Thr Val
195          245          250          255
196    Asp His Leu Ser Lys Tyr Leu Ala Val Arg Leu Ala Leu Glu Glu Leu
197          260          265          270
198    Arg Ser Lys Gly Glu Ser Asn Gln Met Asn Leu Asp Thr Ala Ser Glu
199          275          280          285
200    Lys Gln Tyr Thr Ile Tyr Ile Ala Thr Ala Ser Gly Gln Phe Thr Val
201          290          295          300
202    Leu Asn Gly Ser Phe Ser Leu Glu Leu Val Ser Glu Lys Tyr Trp Lys
203    305          310          315          320
204    Val Asn Lys Pro Met Glu Leu Tyr Tyr Ala Pro Thr Lys Glu His Lys
205          325          330          335
207 (2) INFORMATION FOR SEQ ID NO: 3:
208   (i) SEQUENCE CHARACTERISTICS:
209       (A) LENGTH: 328 base pairs
210       (B) TYPE: nucleic acid
211       (C) STRANDEDNESS: double
212       (D) TOPOLOGY: linear
213   (ii) MOLECULE TYPE: DNA (genomic)
214   (ix) FEATURE:
215       (A) NAME/KEY: CDS
216       (B) LOCATION: 1..327
217   (ix) FEATURE:
218       (A) NAME/KEY: misc feature
219       (B) LOCATION: 1...325
220       (D) OTHER INFORMATION: Partial mouse sequence; nucleotides
221 1-325 show homology to nucleotides 727-1051 of SEQ ID
222 NO: 1 (human).
223   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
224    AGT GGA GCA GAA GAT AAT GGT GAC AGC TCC CAC TGT AGT AAC GCA TCC      48
225    Ser Gly Ala Glu Asp Asn Gly Asp Ser Ser His Cys Ser Asn Ala Ser
226      1          5          10          15
227    ACA CAC AGC AAC CAG GAA GCG GGC CCG AGT AAC AAA CGG ACC AAA ACC      96
228    Thr His Ser Asn Gln Glu Ala Gly Pro Ser Asn Lys Arg Thr Lys Thr
229          20          25          30
230    TCT GAT GAC TCT GGG CTT GAT CTT GAT AAC AAC AAT GCA GGA GTG GCG      144
231    Ser Asp Asp Ser Gly Leu Asp Leu Asp Asn Asn Asn Ala Gly Val Ala
232          35          40          45
233    ATT GAT CCA GTC ATG GAC GGT GCC AGT GAG ATT GAG TTA GTC TTC AGG      192
234    Ile Asp Pro Val Met Asp Gly Ala Ser Glu Ile Glu Leu Val Phe Arg
235          50          55          60
236    CCC CAT CCA ACT CTT ATG GAA AAG GAC GAC AGC GCA CAG ACG AGA TAC      240
237    Pro His Pro Thr Leu Met Glu Lys Asp Asp Ser Ala Gln Thr Arg Tyr
238          65          70          75          80
239    ATA AAG ACT TCA GGC AAT GCC ACT GTT GAT CAC TTA TCC AAG TAT CTG      288
240    Ile Lys Thr Ser Gly Asn Ala Thr Val Asp His Leu Ser Lys Tyr Leu

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/991,888

DATE: 07/25/2002  
TIME: 14:55:39

Input Set : N:\Crf3\RULE60\09991888.raw  
Output Set: N:\CRF3\07252002\I991888.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; Xaa Pos.2,3,5,6,7,8,9,10,11,12,13,14,15,20,23,24,26,27,28,29,30,31  
Seq#:5; Xaa Pos.32,33,34,35,36,38,39

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/991,888

DATE: 07/25/2002

TIME: 14:55:39

Input Set : N:\Crf3\RULE60\09991888.raw

Output Set: N:\CRF3\07252002\I991888.raw

L:22 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:23 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:281 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0  
L:283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:16  
L:285 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:32